IFW

TRANSMITTAL FORM

(to be used for all correspondence after initial filing)

Application Number	10/825,488
Filing Date	April 15, 2004
First Named Inventor	Ganit, et al
Group Art Unit	2171
Examiner Name	Not Yet Assigned
Attorney Docket Number	301560.01

☐ Sent via Express Mail Label No.:

ENCLOSURES *(check all that apply)*

☐ Fee Transmittal Form
  ☐ Fee Attached

☐ Amendment / Reply
  ☐ After Final
  ☐ Affidavits/declaration(s)

☐ Extension of Time Request

☐ Express Abandonment Request

☒ **Information Disclosure Statement with Form PTO/SB/08B (5 pages)**

☐ Response to Notice to File Missing Parts
  ☐ A copy of the Notice to File Missing Parts Under 37 CFR 1.52 or 1.5

☐ Assignment Papers *(for an Application)*

☐ Drawing(s) (sheets)

☐ Declaration
  ☐ Newly Executed (pages)
  ☐ A copy from a prior application (37 CFR 1.63(d)) (pages)

☐ Licensing-related Papers

☐ Petition

☐ Petition to Convert to a Provisional Application

☐ General Power of Attorney (SB80)
  ☐ 37 CFR 3.73(b) Statement

☐ Terminal Disclaimer

☐ Request for Refund

☐ CD, Number of CD(s) __

☐ After Allowance Communication to TC

☐ Appeal Communication to Board of Appeals and Interferences

☐ Appeal Communication to TC
  (Appeal Notice, Brief, Reply Brief)

☐ Proprietary Information

☐ Status Letter

☐ Application Data Sheet

☐ Request for Corrected Filing Receipt

☒ **Return Receipt Postcard**

☒ **Other Enclosure(s)** *(please identify below):*
  ☒ **Copies of (20) Non-Patent References on CD**
  ☒ **Copy of this Transmittal Form**
  ☐
  ☐

CERTIFICATE OF MAILING OR TRANSMISSION
(Under 37 CFR § 1.8(a))

I hereby certify that this correspondence is being:

☒ deposited with the United States Postal Service on the date shown below with sufficient postage as first class mail in an envelope addressed to: Mail Stop Amendment, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450; or

☐ transmitted by facsimile on the date shown below to the USPTO at (703) _____.

06-03-05	*Sherry Smith*
Date	Signature
	Sherry Smith
	Printed Name

Remarks ☒ The Commissioner is hereby authorized to charge any additional fees required, or credit any overpayments, to Deposit Account No. 50-0463 for the above identified patent application.

SIGNATURE OF ATTORNEY OR AGENT

Signature	*Paul Heynssens*	Reg. No.	47,648
Name of Attorney or Agent	Paul B. Heynssens		
Date	6/3/05	Tel.	(425) 707-3913	Facsimile No.	(425) 708-5046

Assignee Name:	MICROSOFT CORPORATION ONE MICROSOFT WAY REDMOND, WA 98052
Customer Number:	22971

Docket No.: 301560.01   PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

First Named Inventor: Ganti, et al   Attorney Docket No.: 301560.01
Application No.: 10/825,488   Group Art Unit: 2171
Filed: 04/15/2004   Examiner:
Customer No.:   22971   Confirmation Number: 8552
Title: Segmentation of Strings into Structured Records

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria VA 22313-1450

STATEMENT FILED PURSUANT TO
THE DUTY OF DISCLOSURE UNDER 37 C.F.R. §§ 1.56, 1.97 AND 1.98

Sir:

Pursuant to the duty of disclosure under 37 C.F.R. §§ 1.56, 1.97 and 1.98, the

Applicant requests consideration of this Information Disclosure Statement.

This Information Disclosure Statement has been filed before the mailing date of a first

Office Action on the merits in the above-identified case. Accordingly, no fee or

certification is required.

The Applicant hereby makes of record in the above-identified application the

information listed on the attached forms PTO/SB/08A and 08B. The order of

presentation of the references should not be construed as an indication of the

importance of the references. Documents cited on the attached forms PTO/SB/08A and

08B are enclosed unless otherwise indicated. It is respectfully requested that:

1. The Examiner consider completely the cited information, along with any other

information, in reaching a determination concerning the patentability of the present

claims;

2. The enclosed forms be signed by the Examiner to evidence that the cited

information has been fully considered by the Patent and Trademark Office during the

examination of this application;

3. The citations for the information be printed on any patent which issues from

this application.

By submitting this Information Disclosure Statement, the Applicant makes no

representation that a search has been performed, of the extent of any search performed,

or that more relevant information does not exist.

By submitting this Information Disclosure Statement, the Applicant makes no

representation that the information cited in the Statement is, or is considered to be,

material to patentability as defined in 37 C.F.R. § 1.56(b).

By submitting this Information Disclosure Statement, the Applicant makes no

representation that the information cited in the Statement is, or is considered to be, in

fact, prior art as defined by 35 U.S.C. § 102.

Notwithstanding any statement by the Applicant, the Applicant urges the

Examiner to form his own conclusion regarding the relevance of the cited information.

An early and favorable action is hereby requested.

Respectfully submitted,

Date: 6/3/05

MICROSOFT CORPORATION

By: _____

Paul B. Heynssens, Reg. No.: 47,648
Microsoft Corporation
Redmond, Washington
Telephone: (425) 707-3913

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(use as many sheets as necessary)

	Complete if Known
Application Number	10/825,488
Filing Date	April 15, 2004
First Named Inventor	Ganti, et al.
Art Unit	2171
Examiner Name	Not Yet Assigned

Sheet	1	of	3	Attorney Docket Number	301560.01

NON PATENT LITERATURE DOCUMENTS			
Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.)., date, page(s), volume–issue number(s), publisher, city and/or country where published.	T[2]
	1	E. MARSH and D. PERZANOWSKI. MUC-7 Evaluation of IE Technology: Overview of Results, Proceedings of the 7th Message Understanding Conference (MUC-7). Morgan Kaufman, 29 April 1998.	
	2	B. ADELBERG. NoDoSE—A Tool for Semi-Automatically Extracting Structured and Semistructured Data from Text Documents. SIGMOD, 1998 Seattle, WA, USA.	
	3	A. ARASU and H. GARCIA-MOLINA. Extracting Structured Data from Web Pages. SIGMOD 2003, June 9-12, 2003, San Diego, CA.	
	4	R. BAUMGARTNER, S. FLESCA, and G. GOTTLOB. Visual Web Information Extraction with Lixto. Proceedings of the 27th VLDB Conference, Roma, Italy, 2001.	
	5	J. BILMES. What HMMs can do. UWEE Technical report, UWEETR-2002-2003, 2002.	
	6	M. E. CALIFF and R. J. MOONEY. Relational Learning of Pattern-Match Rules for Information Extraction. In Working Notes of AAAl Spring Symposium on Applying Machine Learning to Discourse Processing, pages 6—11, Menlo Park, CA, 1998. AAAI Press.	
	7	S. F. CHEN and J. GOODMAN. An Empirical Study of Smoothing Techniques for Language Modeling. In Proceedings of the 34th Annual Meeting of the ACL, pages 310—318, June 1996.	

Examiner Signature		Date Considered	

* EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] Applicant's unique citation designation number (optional) [2] Applicant is to place a check mark here if English language Translation is attached.

Substitute for form 1449/PTO			Complete if Known	
INFORMATION DISCLOSURE STATEMENT BY APPLICANT *(use as many sheets as necessary)*			Application Number	10/825,488
			Filing Date	April 15, 2004
			First Named Inventor	Ganti, et al.
			Art Unit	2171
			Examiner Name	Not Yet Assigned
Sheet **2**	of	**3**	Attorney Docket Number	301560.01

NON PATENT LITERATURE DOCUMENTS			
Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.)., date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
	8	M. COLLINS and Y. SINGER. Unsupervised Models for Named Entity Classification. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, 1999, Pages 100-110.	
	9	V. CRESCENZI, G. MECCA, and P. MERIALDO. RoadRunner: Towards Automatic Data Extraction From Large Web Sites. Proceedings of the 27th VLDB Conference, Roma, Italy, 2001.	
	10	J. DROPPO, L. DENG, and A. ACERO. Evaluation of the Splice on the Aurora 2 and 3 Tasks. Microsoft Research, Pages 29-32.	
	11	S. FINE, Y. SINGER, and N. TISHBY. The Hierarchical Hidden Markov Model: Analysis and Applications. Machine Learning, 32(1):41—62, 1998.	
	12	D. FREITAG and A. MCCALLUM. Information Extraction with HMM Structures Learned by Stochastic Optimization. In AAAI/IAAI, pages 584—589, Copyright 2001.	
	13	R. GRISHMAN. Information Extraction: Techniques and Challenges. In Information Extraction (International Summer School SCIE-97). Springer-Verlag, 1997.	
	14	M. A. HERNANDEZ and S. J. STOLFO. Real-World Data is Dirty: Data Cleansing and the Merge/Purge Problem. Data Mining and Knowledge Discovery, 2(1):9—37, 1998. © 1998 Kluwer Academic Publishers, Boston. Manufactured in The Netherlands.	
	15	C. A. KNOBLOCK, K. LERMAN, S. MINTON, and I. MUSLEA. Accurately and Reliably Extracting Data From the Web: A Machine Learning Approach. IEEE Data Engineering Bulletin, 23(4):33—41, 2000. Copyright 1999 IEEE.	

Examiner Signature		Date Considered	

* EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] Applicant's unique citation designation number (optional) [2] Applicant is to place a check mark here if English language Translation is attached.

NON PATENT LITERATURE DOCUMENTS			
Examiner Initials*	Cite No.[1]	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.)., date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[2]
	16	M. LAPATA. Probabilistic Text Structuring: Experiments with Sentence Ordering. In Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics, July 2003, pp. 545-552.	
	17	A. F. MARTIN and M. A. PRZYBOCKI. NIST 2003 Language Recognition Evaluation. In Eurospeech 2003, 2003.	
	18	A. MIKHEEV, M. MOENS, and C. GROVER. Named Entity Recognition Without Gazetteers. In Proceedings of EACL, 1999.	
	19	L. R. RABINER. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. Proceedings of the IEEE, Vol. 77, No. 2, February 1989. Pages 257-286	
	20	K. SEYMORE, A. MCCALLUM, and R. ROSENFELD. Learning Hidden Markov Model Structure for Information Extraction. In AAAI 99 Workshop on Machine Learning for Information Extraction, 1999.	

Examiner Signature		Date Considered	

* EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] Applicant's unique citation designation number (optional) [2] Applicant is to place a check mark here if English language Translation is attached.